

## Introduction

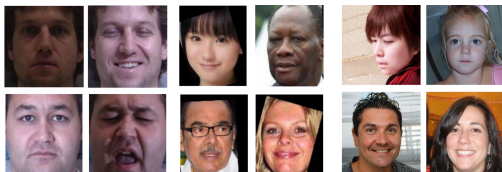
Generative Models are basic building blocks in most of image recognition architectures. GANs can be used to obtain synthetic data where data is scarce and in scenarios where privacy is important. In this work, we analyze bias and fairness of GANs and their impact on face verification.

## Contributions

**Result-1:** We observed that GANs trained on FFHQ dataset exhibit bias for the "age" and "race" protected attributes.

**Result-2:** We demonstrate that Face Verification systems that are trained or fine-tuned with GAN data exacerbate bias for the "race" protected attribute.

## Datasets



(a) CMU Multi-PIE

(b) BFW

(c) FFHQ

**CMU Multi-PIE** - Constrained Face dataset

**BFW** Balanced Faces in the Wild for face verification

**Synthetic Faces** are generated with DiscoFaceGAN

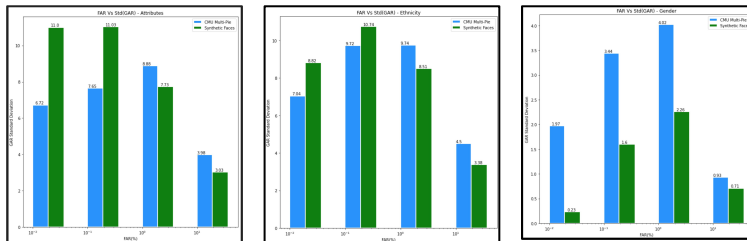
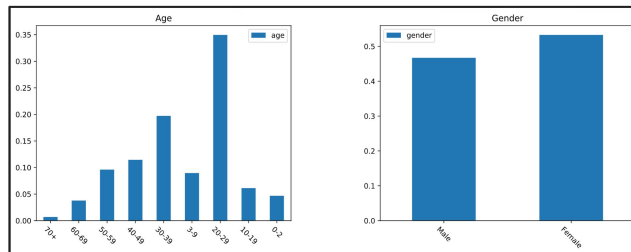


Synthetic(DiscoFaceGAN) Faces

## Experimental Setup

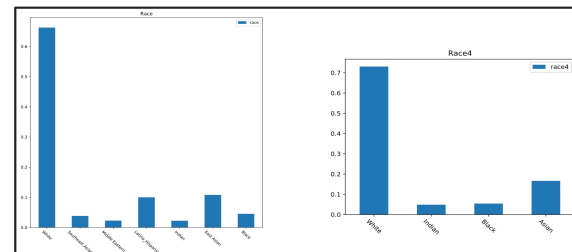
- ❖ **StyleGAN** generated faces are classified with Fairface. Proportion of faces per attribute were plotted
- ❖ **VGGFace2** is fine-tuned with CMU Multi-PIE and Synthetic Faces. Comparison of  $DoB_{FV}$  i.e Std(GAR @ FAR) for different attributes is carried out.

## Results & Observations: Bias & Fairness



## Conclusion

- ❖ Biases present in dataset such as FFHQ are also being manifested in GANs, specifically for attributes such as age, race and gender for faces.
- ❖ The biases could also impact the sub-group performance of downstream models such as face verification systems. Hence, it is important to debias GANs.



- ❑ GANs are biased towards age group "20-29" and "White" faces.
- ❑ Face Verification models trained or fine-tuned with Synthetic faces exhibit bias for "race" attribute.
- ❑ Face Verification models trained or fine-tuned with Synthetic faces doesn't exhibit any bias for "gender" attribute.

- ❑ At, high FAR rates no bias is observed (low  $DoB_{FV}$ ). We hypothesize that although biases are present these are masked by high false acceptances.